Blind deconvolution for thin layered confocal imaging

Praveen Pankajakshan, Bo Zhang, Laure Blanc-Féraud, Zvi Kam,
Jean-Christophe Olivo-Marin, Josiane Zerubia

On blind deconvolution for thin layered confocal imaging

Praveen Pankajakshan,[1,*] Bo Zhang,[2] Laure Blanc-Féraud,[1] Zvi Kam,[3]

Jean-Christophe Olivo-Marin,[2] and Josiane Zerubia[1]

[1] *ARIANA Project-team, INRIA/I3S,*

2004 Route des lucioles, BP 93, 06902 Sophia-Antipolis Cedex, France

[2] *Quantitative Image Analysis Unit,*

Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France

[3] *Department of Molecular Cell Biology,*

Weizmann Institute of Science, Rehovot, Israel 76100.

[*] *Corresponding author: ppankaja@sophia.inria.fr*

In this paper, we have proposed an Alternate Minimization (AM) algorithm for estimating the Point-Spread Function (PSF) of a Confocal Laser Scanning Microscope (CLSM) and the specimen fluorescence distribution. A 3-D separable Gaussian model is used to restrict the PSF solution space and a constraint on the specimen is used so as to favor the stabilization and convergence of the algorithm. The results obtained from the simulation show that the PSF can be estimated to a high degree of accuracy, and those on real data show better deconvolution as compared to a full theoretical PSF model.

OCIS codes: 100.1455, 180.1790, 100.3190.

1. Introduction

Most of the fluorescence microscopes that image a uniformly illuminated three-dimensional (3-D) object by the optical sectioning technique, are affected by some out-of-focus fluorescence contributions. Secondary fluorescence from the sections away from the region of interest often interferes with the contrast and resolution of those features that are in focus. Let us take the case of a single-photon (1-p) fluorescence microscope like the Widefield Microscope (WFM) and the Confocal Laser Scanning Microscope (CLSM) [1]. For the sake of simplicity, if we assume that the detectors are the same, then a WFM could be seen as a CLSM but with a fully-open pinhole. The WFM can collect more light even from the deeper sections of a specimen but the

data are sometimes rendered useless as there is a significant amount of out-of-focus blur. The maximum intensity in each plane decreases as z^{-2}, with z being the axial distance from the source. A completely closed pinhole (diameter < 1 Airy Units (AU); 1AU$= 1.22\lambda_{ex}$/Numerical Aperture) on the other hand, confines the light detected only to the in-focus plane but at the expense of imaging low-contrast, highly noisy (signal dependant noise) images. The intensity from a point source in this case decreases as z^{-4} and the loss of in-focus intensity inhibits imaging of weakly fluorescent specimens. Even with a useable pinhole diameter of 1AU, 30% of the light collected is from the out-of-focus regions. In addition, the microscope is inherently *diffraction-limited* [1, 2] and the image of a point source (the Point-Spread Function (PSF)) displays a lateral diffractive ring pattern (expanding with defocus) introduced by the finite-lens aperture.

Let $\mathcal{O}(\Omega) = \{o = (o_{xyz}) : \Omega \subset \mathbb{N}^3 \to \mathbb{R}\}$ denote all possible observable objects on the discrete spatial domain $\Omega = \{(x, y, z) : 0 \leq x \leq N_x - 1, 0 \leq y \leq N_y - 1, 0 \leq z \leq N_z - 1\}$ and $h : \Omega \mapsto \mathbb{R}$ the microscope PSF. If we assume that the imaging system is linear and shift-invariant, then the interaction between h and o is a "3-D convolution": $(h * o)(\mathbf{x}) = \sum_{\mathbf{x}' \in \Omega} h(\mathbf{x} - \mathbf{x}')o(\mathbf{x}')$. From the perspective of computational methods, this could be inverted with the knowledge of the scanning system properties and also by information about the object being scanned. It is for this reason that the knowledge of the PSF h is of fundamental importance. The nature of the PSF for fluorescence microscope has been studied extensively [3–5]. We will introduce the

reader in Section 2.B to one such theoretical model based on the scalar diffraction theory and its parametric approximation in Section 3.B.

1.A. Problem Formulation

Restoration by deconvolution could be achieved by using either a *non-blind* or *blind* approach. For the non-blind case, the most common approach is an experimental procedure [6, 7] that obtains the PSF by imaging a small fluorescent bead (so as to approximate a point object) positioned in the cover slide. Although such a PSF should have been an ideal choice for a deconvolution algorithm, it suffers from low contrast (can be recorded only at finite defocus ranges) and is contaminated by noise. A way to suppress the noise would be to either acquire several bead data sets and then average them [8, 9] or reconstruct them using Zernike polynomial moments [10]. This approach is however handicapped by alignment problems and also the whole process could take a long time. The alternative would be to use an analytical model of the PSF [11, 12] that takes into account the acquisition system's physical information as parameters. This information however might not be available or might change during the course of the experiment (for example, due to heating of live samples).

We hence arrive at the blind deconvolution approach of estimating the specimen and the unknown PSF parameters using a single observation of the specimen volume. The problem of blind deconvolution is thus reduced to answering the following question:

"*How does one estimate the original object and the PSF, given only a single observation?*"

If we forget the effect of noise and consider the observation model $(h * o)$ in the Fourier space as: $\mathcal{F}(i) = \mathcal{F}(h) \cdot \mathcal{F}(o)$, several solutions for o and h answer this problem. For example, if (h, o) is a solution, then the trivial case is that h is a Dirac function and $o = i$ or vice versa. If h is not irreducible, there exists h_1 and h_2 such that $h = h_1 * h_2$, and the couples ($h_1 * h_2$, o) and (h_1, $h_2 * o$) are also solutions. Another ambiguity is in the scaling factor. If (h, o) is a solution, then $(\tau h, 1/\tau o)$ $\forall \tau > 0$ are solutions too. This last ambiguity can be waived for example by imposing a forced normalization on h. Thus broadly speaking, a way of reducing the space of possible solutions and to regularize the problem is to introduce constraints on h and o. If the problem of deconvolution is ill-posed, that of blind deconvolution is under-determined as the number of unknowns to be estimated is increased without any increase in the input observation data.

Many methods use an iterative approach to estimate the PSF and the object with no prior information on the object [13, 14]. Markham and Conchello [15] worked on a parametric form for the PSF and developed an estimation method utilizing this model. The difficulty in using this model for our application is that the number of free parameters to estimate is large and the algorithm is computationally expensive. Hom *et al.* [16] proposed a myopic deconvolution algorithm that alternates between iteration to deconvolve the object and estimate the PSF. In order to myopically recon-

struct the PSF, they introduce a constraint on the Optical Transfer Function (OTF) (the OTF and the PSF are Fourier Transform pairs).

This paper is organized in the following manner: we first discuss the nature of the noise, its mathematical modeling and handling in Section 2.A. The PSF modeling is introduced in the Section 2.B. Section 3 is dedicated to the proposed joint restoration and estimation of the imaged object and the microscope PSF using a Bayesian framework. Direct restoration from the observation data is very difficult, and hence it is necessary to define an underlying model for both the object and the PSF respectively. An Alternate Minimization (AM) algorithm is then proposed to solve this particular problem. This AM algorithm was then tested on images of degraded phantom objects and real data; the results obtained are presented in Section 4. We then conclude in Section 5 with a discussion and proposed future work. The scope of this paper is restricted to restoring images from a CLSM given the spatial invariance nature of the diffraction-limited PSF.

2. Sources of distortion and their modeling

2.A. Poissonian Assumption

In digital microscopy, the source of noise is either the signal itself (so-called 'photon shot noise'), or the digital imaging system. By tracking the photon to electron conversion at the detector, we can observed that the signal and the dependent noise follows an underlying distribution which is Poissonian [17]. Conversely, the imaging

noise isolated in the absence of any fluorescence source follows a Gaussian distribution [18,19]. The interested reader may refer to [1,20] for more details on this subject.

In this paper, we have assumed that there is no *readout* or *dark* noise as the Photomultiplier Tube is operating in the photon-counting mode. When the imaging system has been a priori calibrated, there is almost negligible offset in the detector and the illumination is uniform. Thus if $\{i(\mathbf{x}) : \mathbf{x} \in \Omega\}$ (assumed to be bounded and positive) denotes the observed intensity of the volume, for the Poissonian assumption, the observation model can be expressed as:

$$\gamma i(\mathbf{x}) = \mathcal{P}(\gamma([h * o](\mathbf{x}) + b(\mathbf{x}))), \mathbf{x} \in \Omega, \tag{1}$$

where, $\mathcal{P}(\cdot)$ denotes voxel-wise noise function modeled as a Poissonian process. $b : \Omega \mapsto \mathbb{R}$ is a uniformly distributed intensity that models the low-frequency background signal caused by scattered photons and auto-fluorescence from the sample. $1/\gamma$ is known as the photon conversion factor, and $\gamma i(\mathbf{x})$ is the observed photon at the detector.

2.B. Theoretical diffraction-limited PSF Model

Among the enormous literature available on PSF modeling, we highlight the work of P. A. Stokseth [11] who obtained the OTF for an aberration-free optical system especially for large defocus. This model was used to study the PSFs under different microscope settings, and also in validating the algorithm.

If we consider a converging spherical wave in the object space from the objec-

tive lens, the near-focus amplitude distribution h_A can be written in terms of the amplitude optical transfer function OTF_A as: $h_A(\mathbf{x}) = \int_{\mathbf{k}} OTF_A(\mathbf{k}) exp(j\mathbf{k}\cdot\mathbf{x}) d\mathbf{k}$, where, $j^2 = -1$, and $\mathbf{x}$ and $\mathbf{k}$ are the 3-D coordinates in the image and the Fourier space respectively. By making the axial Fourier space coordinate k_z a function of lateral coordinates, $k_z = (k^2 - (k_x^2 + k_y^2))^{1/2}$, the 3-D Fourier transform is reduced to: $h_A(x, y, z) := \int_{k_x} \int_{k_y} P(k_x, k_y, z) exp(j(k_x x + k_y y)) dk_y dk_x$, where, $k = 2\pi\mu/\lambda$ is the wave number of an illumination wave with a wavelength λ in vacuum and in a medium of refractive index μ, and $P(\cdot, \cdot)$ describes the overall complex field distribution in the pupil of an non-aberrated objective lens [2, 11].

For an aberration-free microscope, the pupil function can be written as:

$$P(k_x, k_y, z) = \begin{cases} A(\phi) exp(jk\psi), & \text{if } \sqrt{k_x^2 + k_y^2} < k \sin\phi_{max} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where, ψ is the optical distance between the wavefront emerging from the exit pupil and the reference sphere measured along the extreme ray, $\phi = \sin^{-1}(k_x^2 + k_y^2)^{1/2}/k$, and ϕ_{max} is the maximum semi-aperture angle of the objective. The intensity projected from an isotropically illuminating point source such as a fluorophore, on a (flat) pupil plane is bound to be energy conservation constraint. Therefore, the amplitude $A(\phi)$ in the pupil plane for detection should vary as $(\cos\phi)^{-1/2}$ and the energy as $(\cos\phi)^{-1}$ [21]. Conversely, for the illumination case, $A(\phi)$ varies as $(\cos\phi)^{1/2}$. Also for small defocus, ψ in (2) could be approximated as [11]: $\psi = z(1 - \cos\phi)$. To derive the

intensity distribution of a point source in the image space of a CLSM, we make use of the Helmholtz reciprocity theorem. Since in induced fluorescence, the excitation (λ_{ex}) and the emission wavelengths (λ_{em}) are different, the confocal PSF can be written as [22]:

$$h(\mathbf{x}) = C|h_A(\mathbf{x};\lambda_{ex})|^2 \cdot \int\limits_{x_1^2+y_1^2 \leq \frac{D^2}{4}} |h_A(x-x_1, y-y_1, z;\lambda_{em})|^2 dx_1 dy_1, \tag{3}$$

where C is a scaling factor and D is the back-projected diameter of the circular pinhole. This theoretical model of the PSF does not take into account aberrations and assumes that diffraction effect predominates the aberrations. However, this scalar model could be extended for other aberrations by modifying the pupil function expression in (2) to also include the additional phase term due to aberrations [20].

3. Bayesian framework for the Alternate Minimization (AM) blind deconvolution algorithm

In this section we will use the Bayesian framework to describe the method for the blind deconvolution.

3.A. Deconvolution

Since the advent of the nearest neighbor deconvolution algorithm [3], there have been numerous techniques proposed [23–26] for image restoration applied to microscopy. These however assume that the noise is Gaussian and are valid only for images with high SNR. Statistical methods [27,28] on the other hand are extremely effective when

the noise in the acquired 3-D image is fairly strong. We propose here one such non-linear iterative algorithm which although slightly computationally expensive (in comparison to linear methods), can better restore the lost higher frequencies.

If we accept the Poissonian model approximation of (1), then the image i can be interpreted as the realization of independent Poison processes at each voxel. Hence the likelihood can be written as:

$$\Pr(i|o,h) = \prod_{\mathbf{x}\in\Omega} \frac{[h*o](\mathbf{x})^{i(\mathbf{x})} e^{-[h*o](\mathbf{x})}}{i(\mathbf{x})!}, \tag{4}$$

where, the mean of the Poisson process is given by $[h*o](\mathbf{x})$. In all the derivations used henceforth, the background term has been excluded but the algorithm can be modified by changing the above mean to $[h*o+b](\mathbf{x})$. The background fluorescence can be determined from the smoothed histogram of a single "specimen-independent" slice, and it is subsequently added to the mean at every iteration of the Maximum Likelihood (ML) algorithm (4) for o [20]. As iterative ML methods do not ensure any smoothness constraints, if unchecked, they evolve to a solution that displays many artifacts from noise amplification (for examples see [29]). There are many remedies like terminating the iteration (manually or by using a statistical criterion) before the deterioration begins or pre-filtering the observation data. One might argue that by applying a low-pass filter as a pre-processing step before deconvolution (as in [30]), the results are improved in comparison to the deconvolved images with no pre-filtering. The deconvolution algorithm applied after denoising is less influenced by the

prior term of the object [31]. However, such pre-filtering operations might influence the blind deconvolution algorithm as it is not clear how the resulting filtered data is eventually mapped to the original object. The number of iterations for eventual convergence of the deconvolution algorithm also increases and the final result need not be optimum. Such interventions are thus a post hoc method of regularizing the ill-posed problem as it is a way of bringing some knowledge about the solution o. The Maximum a Posteriori (MAP) algorithm proposed in this paper uses the prior model on the specimen and the PSF but within the Bayesian framework. We are hence able to simultaneously denoise and deconvolve the observation data without making any modifications whatsoever.

By using the Bayes theorem and assuming that o and h are independent, the posterior joint probability is:

$$\Pr(o, h|i) = \frac{\Pr(i|o, h)\Pr(o)\Pr(h)}{\Pr(i)} \tag{5}$$

where, $\Pr(o)$ is the global prior probability on the object and $\Pr(h)$ is the global prior on the PSF. The nature of the prior terms and their expressions are discussed in the Sections 3.A.1 and 3.B. The estimates for o and h can be obtained by simultaneously maximizing the joint probability as:

$$\begin{aligned}(\hat{o}, \hat{h}) &= \arg\max_{(o,h)}\{\Pr(o, h|i)\} \\ &= \arg\min_{(o,h)}\{-\log[\Pr(o, h|i)]\}. \end{aligned} \tag{6}$$

As $\Pr(i)$ does not depend on o or h, it shall hereafter be excluded from all the estimation procedures that involve either o or h. The minimization of the cologarithm of $\Pr(o, h|i)$ in (6) can be rewritten as the minimization of the following energy functional:

$$\mathcal{J}(o, h|i) \equiv \mathcal{J}_{obs}(i|o, h) + (\lambda_o \mathcal{J}_{reg,o}(o) + \lambda_h \mathcal{J}_{reg,h}(h)). \tag{7}$$

$\mathcal{J}_{obs} \colon \Omega \mapsto \mathbb{R}$ is a measure of fidelity to the data and it corresponds to the term $\Pr(i|o, h)$ which is given from the noise distribution. It has the role of pulling the solution towards the observed data. While, $\mathcal{J}_{reg,o} \colon \Omega \mapsto \mathbb{R}$ and $\mathcal{J}_{reg,h} \colon \Omega \mapsto \mathbb{R}$ are the prior terms on the object and the PSF which ensure smoothness of the solutions. λ_o and λ_h are positive parameters which measures the trade off between goodness of fit and the regularity of the solutions. For the Bayesian interpretation of regularization problems, we refer the reader to the paper [32] by Demoment.

Practically, simultaneous estimation of o and h from (6) is a difficult task. A way to overcome this difficulty is to alternatively maximize the posterior first with respect to o while assuming that the PSF function h is known and fixed, and then update the PSF using the previous object estimate. This joint optimization algorithm is summarized as:

$$\begin{aligned} \hat{o}^{(n+1)} &= \arg\max_{o} \{\Pr(i|o, \hat{h}^{(n)}) \Pr(o)\}, \\ \hat{h}^{(n+1)} &= \arg\max_{h} \{\Pr(i|\hat{o}^{(n+1)}, h) \Pr(h)\}. \end{aligned} \tag{8}$$

The implementation strategy of this blind deconvolution schema has been shown in Algorithm 1 on Section 4.A and the discussion follows in the subsequent Sections.

3.A.1. A priori object models

The ensemble model of an object class refers to any probability distribution $\Pr(o)$ on the object space $\mathcal{O}$ of the following form:

$$\Pr(o) = Z_{\lambda_o}^{-1} e^{-\lambda_o E(o)}, \tag{9}$$

where, $E(o)$ is a generalized energy and $1/\lambda_o$ (with $\lambda_o > 0$) is the Gibbs parameter for the prior term. We associate with each site $(x, y, z) \in \Omega$ of the object a unique neighborhood $\eta_{xyz} \subseteq \Omega \setminus (x, y, z)$, and we denote the collection of all neighbors $\eta = \{\eta_{xyz} | (x, y, z) \in \Omega\}$ as the neighborhood system. If we assume that the random field $(O = o)$ on a domain Ω is Markovian with respect to the neighborhood system η, then, $\Pr(o_{xyz} | o_{\Omega \setminus x,y,z}) = \Pr(o_{xyz} | o_{\eta_{xyz}})$. o is a Markov Random Field (MRF) on (Ω, η), if o denotes a Gibbs ensemble on Ω and the energy is a superposition of potentials associated to the cliques (a set of connected pixels). Hence, $E(o) = \sum_{C \in \mathcal{C}} V_C(o)$.

We use in this paper, the following first-order, homogeneous, isotropic MRF, over a 6 member neighborhood $\eta_{\mathbf{x}} \in \eta$ (see Fig. 1) of the site $\mathbf{x} \in \Omega$,

$$\Pr[O = o(\mathbf{x})] = Z_{\lambda_o}^{-1} e^{-\lambda_o \sum\limits_{\mathbf{x}} |\nabla o(\mathbf{x})|} \tag{10}$$

where, $|\nabla o(\mathbf{x})|$ is the *potential function* and λ_o is the regularization parameter described above. The estimation of this parameter is dealt with in Section 4.A.

From a numerical point of view, $|\nabla o(\mathbf{x})|$ is not straightforward to minimize, since it is not differentiable in zero. An approach to circumvent this problem is to regularize it, and instead to consider the (isotropic) discrete definition as:

$$\begin{aligned} |\nabla o(x,y,z)|_\epsilon &= ((o(x+1,y,z)-o(x,y,z))^2+(o(x,y+1,z)-o(x,y,z))^2 \\ &\quad +(o(x,y,z+1)-o(x,y,z))^2+\epsilon^2)^{\frac{1}{2}}, \end{aligned} \tag{11}$$

where, ϵ is an arbitrarily small value ($< 10^{-3}$). For the *partition function* $Z_{\lambda_o} = \sum_{o\in\mathcal{O}(\Omega)} exp(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|_\epsilon)$ to be finite, we restrict the possible values of $o(\mathbf{x})$ so that the numerical gradient of $\nabla o(\mathbf{x})$ is also bounded. When this model is used as a prior for the object, we have the following smoothed regularization functional:

$$\mathcal{J}_{reg,o}(o(\mathbf{x})) = \lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|_\epsilon. \tag{12}$$

For numerical calculations, we will use the above smoothed approximation, and $|\nabla o(\mathbf{x})|_\epsilon$ will henceforth be simply written as $|\nabla o(\mathbf{x})|$. From (10) and (12), it can be inferred that sites with very high total gradient intensities are more penalized and those with low total gradients are less penalized. This is because it is more likely that high gradients correspond to the case where there is less similarity between the site of interest and their closest neighbors.

Tikhonov-Miller [33,34] introduced a regularization based on the ℓ_2 norm of the gradient of the image. However, we have used the Total Variation (TV) regularization (see Rudin-Osher-Fatemi (ROF) [35]) as it is able to better preserve discontinuities [28]. A direct 3-D extension of this TV algorithm for CLSM is described in [29] and [36].

3.A.2. Estimation of the object

For the time being let us assume that either the PSF or its parameters $\boldsymbol{\theta} \in \Theta$ are known (either by initialization or from previous estimates) and hence $\hat{h}$ is determinate. From (4), (5) and (10) we get:

$$\Pr(o, \hat{h}|i) = Z_{\lambda_o}^{-1} e^{-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|} \cdot \prod_{\mathbf{x}\in\Omega} \frac{[\hat{h} * o](\mathbf{x})^{i(\mathbf{x})} e^{-[\hat{h}*o](\mathbf{x})}}{i(\mathbf{x})!}. \tag{13}$$

As in (6), by applying $-log$ operator to the a posteriori above, the cost function $\mathcal{J}(o, \hat{h}|i)$ to be minimized with respect to o becomes:

$$\mathcal{J}(o, \hat{h}|i) \equiv \Big(\sum_{\mathbf{x}\in\Omega} [\hat{h} * o](\mathbf{x}) - \sum_{\mathbf{x}\in\Omega} i(\mathbf{x}) \log[\hat{h} * o](\mathbf{x}) + \sum_{\mathbf{x}\in\Omega} \log(i(\mathbf{x})!)\Big) + \lambda_o \sum_{\mathbf{x}\in\Omega} |\nabla o(\mathbf{x})| + \log[Z_{\lambda_o}]. \tag{14}$$

Richardson-Lucy (RL) algorithm with TV regularization The Euler-Lagrange equation for minimizing $\mathcal{J}(o, \hat{h}|i)$ in (14) with respect to o is:

$$1 - \hat{h}(-\mathbf{x}) * \Big(\frac{i(\mathbf{x})}{(\hat{h} * o)(\mathbf{x})}\Big) - \lambda_o \, div\Big(\frac{\nabla o(\mathbf{x})}{|\nabla o(\mathbf{x})|}\Big) = 0 \tag{15}$$

where, $\hat{h}(-\mathbf{x})$ is the Hermitian adjoint operation on $\hat{h}(\mathbf{x})$ and div stands for the divergence (see [28] for details). Inspired by the RL algorithm [37, 38], (15) can be solved for the object o by the following fixed-point iterative algorithm:

$$\hat{o}^{(n+1)}(\mathbf{x}) = \left[\frac{i(\mathbf{x})}{(\hat{o}^{(n)} * \hat{h})(\mathbf{x})} * \hat{h}(-\mathbf{x})\right] \cdot \frac{\hat{o}^{(n)}(\mathbf{x})}{1 - \lambda_o \, div\Big(\frac{\nabla \hat{o}^{(n)}(\mathbf{x})}{|\nabla \hat{o}^{(n)}(\mathbf{x})|}\Big)} \tag{16}$$

where, $(\cdot)$ denotes the Hadamard multiplication (component wise) and n the iteration number for the deconvolution algorithm. (16) is similar to the Expectation-

Maximization (EM) algorithm [39] with an underlying statistical model of the process, and can be used for obtaining the MAP estimate of the object. The term $div\Big(\nabla\hat{o}^{(n)}(\mathbf{x})/|\nabla\hat{o}^{(n)}(\mathbf{x})|\Big)$ can be numerically implemented with the use of central-differences and the *minmod* scheme [29].

Positivity and flux constraint for the object estimate The deconvolution algorithm that was described above suffers from an inherent weakness. For large values of λ_o, even when the starting guess $\hat{o}^{(n)}$ (with $n = 0$) is positive, the successive estimates need not necessarily have positive intensities. We know that the true intensity of the object $o(\mathbf{x})$ is always non-negative. Most algorithms often truncate these negative intensities to zero or a small positive value. This however is a crude manner to handle the estimated intensities as it can lead to loss of some essential information and sometimes also introduce bias into the calculations.

So how else can the problems associated with negative intensity estimates be handled? Fortunately, the problem is entirely due to poor statistical methodology. The modification that we suggest is to include this knowledge of non-negative true intensities into the prior term of (10). The distribution that would express precisely this condition is:

$$\Pr[o(\mathbf{x})] := \begin{cases} Z_{\lambda_o}^{-1} e^{-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|}, & \text{if } o(\mathbf{x}) \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

For the sake of numerical differentiability we approximate the prior (17) using a *sigmoid* function as:

$$\Pr[o(\mathbf{x})] := Z_{new,\lambda_o}^{-1} e^{-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|} \cdot \left(\frac{1}{(1+\exp(\beta_o(\epsilon - o(\mathbf{x}))))} \right) \tag{18}$$

where, ϵ is a small value close to zero, and β_o is a value that specifies the steepness of the *sigmoid* curve. Typically the values of β_o and ϵ are chosen to be very large and small respectively as precision allows. Their values do not individually affect the algorithm and hence need not be known accurately. The cost function (14), the Euler-Lagrange equation (15) and the multiplicative algorithm (16) are thus modified as follows:

$$\mathcal{J}(o, \hat{h}|i) \equiv \Big(\sum_{\mathbf{x}\in\Omega} [\hat{h} * o](\mathbf{x}) - \sum_{\mathbf{x}\in\Omega} i(\mathbf{x}) \log[\hat{h} * o](\mathbf{x}) \Big) + \lambda_o \sum_{\mathbf{x}\in\Omega} |\nabla o(\mathbf{x})| + \log[Z_{new,\lambda_o}] - \log\left(\frac{1}{(1+\exp(\beta_o(\epsilon - o(\mathbf{x}))))} \right), \tag{19}$$

$$1 - \hat{h}(-\mathbf{x}) \Big(\frac{i(\mathbf{x})}{(h * o)(\mathbf{x})} \Big) - \lambda_o \, div \Big(\frac{\nabla o(\mathbf{x})}{|\nabla o(\mathbf{x})|} \Big) - \beta_o \frac{\exp(\beta_o(\epsilon - o(\mathbf{x})))}{1+\exp(\beta_o(\epsilon - o(\mathbf{x})))} = 0, \tag{20}$$

$$\hat{o}^{(n+1)}(\mathbf{x}) = \left[\frac{i(\mathbf{x})}{(\hat{o}^{(n)} * \hat{h})(\mathbf{x})} * \hat{h}(-\mathbf{x}) \right] \cdot \frac{\hat{o}^{(n)}(\mathbf{x})}{1 - \lambda_o \, div \Big(\frac{\nabla \hat{o}^{(n)}(\mathbf{x})}{|\nabla \hat{o}^{(n)}(\mathbf{x})|} \Big) - \beta_o \frac{\exp(\beta_o(\epsilon - o(\mathbf{x})))}{1+\exp(\beta_o(\epsilon - o(\mathbf{x})))}}. \tag{21}$$

Intuitively, the cost function in (19) ensures that the energy for negative intensity pixels ($o(\mathbf{x}) < \epsilon$) is very high and hence is not reachable (or is not a possible solution) during the iteration procedure.

If PSF is normalized such that $||\hat{h}(\mathbf{x})||_1 = 1$, in the absence of a background signal, it is simple to show that for each iteration of the RL algorithm (see (16) with $\lambda_o = 0$), the following property is true: $s = \sum_{\mathbf{x}\in\Omega} i(\mathbf{x}) = \sum_{\mathbf{x}\in\Omega} \hat{o}(\mathbf{x})$. This property is known as the flux or global photometry conservation and it guarantees that the total number of counts of the reconstructed object is the same as the total number of observation counts. However, this property is lost with regularization and can be incorporated by modifying the cost function (14) to an additive form or by enforcing it in the following manner after every iteration: $\hat{o}^{(n+1)}_{\text{new}}(\mathbf{x}) = (s^{(0)} \times \hat{o}^{(n+1)}_{\text{old}}(\mathbf{x}))/s^{(n+1)}$, where, $s^{(n+1)} = \sum_{\mathbf{x}\in\Omega} \hat{o}^{(n+1)}_{\text{old}}(\mathbf{x})$, and $s^{(0)} = \sum_{\mathbf{x}\in\Omega} \hat{o}^{(0)}(\mathbf{x}) = \sum_{\mathbf{x}\in\Omega} i(\mathbf{x})$.

3.B. Parametrization of the Point-Spread Function

When $\lambda_o = 0$ in (14), theoretically speaking the estimation method on the object and PSF should be the same as h and o play a symmetric role. When no constraint is imposed on the PSF, the solution is not always unique. Some reason that a regularization model on the PSF ($\mathcal{J}_{reg,h}(h)$) could also be argued along the same lines as the constraints introduced earlier for o [14, 40]. Firstly, a Total Variation (TV) [35] kind of regularization cannot model the continuity and regularity in the PSF. A ℓ_1 kind of norm is suitable only for PSFs that have edges, like motion blur [41]. Secondly, in such cases the recovered PSF will be very much dependent on the object/specimen [42]. Separation of the PSF and the object in this case becomes difficult as they have the same or similar solution space. Finally, the regularization parameter λ_h for such a

model is highly dependent on the amount of defocus, and varies drastically from one image sample to another. It is for these reasons that we are proposing to intrinsically regularize the PSF through a parametric model.

Due to the invariance property of ML estimation, we can say: $\hat{h}_{ML}(\mathbf{x}) = h(\mathbf{x}; \hat{\boldsymbol{\theta}}_{ML})$ is the MLE of the PSF. $\theta \in \Theta \subset \mathbb{R}^+$ is the set of parameters that defines the PSF. In a more general manner, any PSF can be written as the decomposition on a set of basis functions $\boldsymbol{\Phi}$ as: $h(\mathbf{x}) \approx \sum_{l=1}^{N_b} w_l \Phi_l(\mathbf{x}) = \langle \mathbf{w}, \boldsymbol{\Phi}(\mathbf{x}) \rangle, \forall \mathbf{x} \in \Omega$, where, w_l denotes the corresponding weights, and N_b denotes the number of the basis functions. The imperfections in an image formation system normally act as passive operations on the data, i.e. they neither absorb nor generate energy. Thus, when an object goes out of focus it is blurred, but the volume's total intensity remains constant. Consequently, all energy arising from a specific point in the fluorescent specimen should be preserved and $\|h(\mathbf{x})\|_1 = \sum_{\mathbf{x} \in \Omega} |h(\mathbf{x})| = 1$. From (3), it is clear that the intensity distribution of a point source will always be positive and so $h(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \Omega$. To satisfy the above defined conditions, and an additional criterion of circular symmetry (i.e. $h(-x, -y) = h(x, y), \forall (x, y) \in \mathbb{R}^2$), the Gaussian kernel is chosen as the basis (see [43] for the 2-D case). This drastically reduces the number of free parameters to estimate and yet retain a reasonable fit to the actual PSF. It was demonstrated by Zhang et al. [5] that for a CLSM, a 3-D separable Gaussian model gives a Relative Squared Error (RSE)$< 9\%$ for a pinhole diameter $D < 3$AU and when the PSF peaks being matched (i.e. $\|h(\mathbf{x})\|_\infty = 1$). Where we say $RSE := \|PSF - h\|_2^2 / \|PSF\|_2^2$.

Thus the diffraction-limited PSF (with restrictions on the pinhole diameter D) can be approximated as:

$$h(\mathbf{x}) = (2\pi)^{-\frac{3}{2}}|\Sigma|^{-\frac{1}{2}}exp(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})). \tag{22}$$

where, $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)^T$ is the mean vector, $\Sigma = [\sigma_{ij}]_{1\leq i,j\leq 3}$ is the covariance matrix, and $\mu_{(\cdot)}$, $\sigma_{(\cdot,\cdot)} \in \Theta$. As a first approximation, for thin-layered specimen imaging with no aberrations, the PSF is spatially invariant and $\boldsymbol{\mu} = \{\mathbf{0}\}$. A mirror symmetry about the central xy-plane results in a diagonal covariance matrix and hence its determinant is $|\Sigma| = \sigma_r^4\sigma_z^2$, where $\sigma_{11}(=\sigma_{22}) = \sigma_r$ and $\sigma_{33} = \sigma_z$ are the lateral and axial spreads respectively. It can be shown that the parameters that we are interested in estimating $\boldsymbol{\theta} = \{\sigma_r, \sigma_z\}$, are dependant on the following settings: wavelength λ_{ex}, refractive index μ and the numerical aperture NA [5].

PSF parameter estimation on the complete data

The method outlined in Section 3.A.2 requires the knowledge of the PSF $\hat{h}(\mathbf{x})$ or $h(\mathbf{x};\hat{\boldsymbol{\theta}})$. From (4), (6), (8), and with the invariance property of ML estimation described earlier, minimizing the energy function with respect to the PSF ($\mathcal{J}(\hat{o},\, h|i)$) or the parameters ($\mathcal{J}(\hat{o},\, \boldsymbol{\theta}|i)$) are equivalent. Thus:

$$\mathcal{J}(\hat{\boldsymbol{o}},\, \boldsymbol{\theta}|i) = -\sum_{\mathbf{x}\in\Omega}(i(\mathbf{x})\log[h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x})) + \sum_{\mathbf{x}\in\Omega}[h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x}). \tag{23}$$

If the true object o is assumed to be known a priori as $\hat{o}$, then estimation of the true parameters of the PSF is straight forward as the cost function (23) is convex in the

neighborhood of optimal $\boldsymbol{\theta} \in \Theta$ (see Fig. 2). The parameters of the PSF can hence be obtained by a Gradient-Descent (GD) kind of algorithm [44]. Analytically minimizing (23) with respect to the parameters leads us to the following:

$$\hat{\boldsymbol{\theta}}_l^{(n+1)} = \hat{\boldsymbol{\theta}}_l^{(n)} - \alpha^{(n)} \nabla_{\boldsymbol{\theta}_l} \mathcal{J}(\hat{o}, \hat{\boldsymbol{\theta}}_l^{(n)} | i);\ \hat{\boldsymbol{\theta}} = \{\hat{\sigma_r}, \hat{\sigma_z}\}, \tag{24}$$

where, $\alpha^{(n)}$ and $\nabla_{\boldsymbol{\theta}_l} \mathcal{J}(\hat{o},\ \hat{\boldsymbol{\theta}}_l^{(n)} | i)$ are the step size and the search direction at iteration n. The gradient of the cost function with respect to the parameters can be calculated as:

$$\nabla_{\boldsymbol{\theta}_l} \mathcal{J}(\hat{o},\ \boldsymbol{\theta}_l | i) = \sum_{\mathbf{x} \in \Omega} [\frac{\partial}{\partial \boldsymbol{\theta}_l} h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x}) - \sum_{\mathbf{x} \in \Omega} \frac{i(\mathbf{x})}{[h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x})} \cdot [\frac{\partial}{\partial \boldsymbol{\theta}_l} h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x}), \tag{25}$$

$\forall \boldsymbol{\theta}_l > 0 \in \Theta$.

If we assume that the PSF is axially and radially centered i.e $\boldsymbol{\mu} = \mathbf{0}$, $\Big[\partial h(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}_l\Big]_{\boldsymbol{\theta}_l = \sigma_r} = \Big(-2/\sigma_r + r^2/\sigma_r^3\Big) h(\boldsymbol{\theta})$, and $\Big[\partial h(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}_l\Big]_{\boldsymbol{\theta}_l = \sigma_z} = \Big(-1/\sigma_z + z^2/\sigma_z^3\Big) h(\boldsymbol{\theta})$. The separable nature of the Gaussian distribution reduces the complexity of the algorithm, as the convolution with the 3-D Gaussian PSF can be implemented as three successive $1D$ multiplications in the Fourier domain. Only a single FFT of the object estimate $\hat{o}$ needs to be performed as an analytical closed form expression for the Fourier transform of the Gaussian and its derivative exists and can hence be numerically calculated. We stop the computation if the difference measure between two successive iterations is smaller than ϵ (in practice 10^{-3} or 10^{-4}), and use the last estimate as the best one.

4. Results

In this section, we validate the proposed AM algorithm on some synthetic and real data.

4.A. Algorithm analysis

The global procedure alternatively minimizes the cost function (14) first with respect to o (16) while keeping the PSF function h fixed and then update the PSF (24) using the previous object estimate $\hat{o}$. Since the iterative algorithm requires an initial guess for the true object, we use the mean of the observed image (i.e. every site is assumed to have a *uniform intensity* and is hence equally likely) for the initialization. For the PSF, as there are no constraints on its spread or support, initialization of the parameters to small values cannot guarantee its convergence to the desired size (due to the Dirac trivial solution). To avoid this problem, we choose the initial parameters to be utmost $2\kappa^{-1}$ Resels and $6\kappa^{-1}$ Resels (1 Resel$= 0.61\lambda_{ex}/NA$; NA: Numerical Aperture; $\kappa = 2.35$) for the lateral and axial case respectively, and descend down to the optimal value. Both $\mathcal{J}_{obs}(i|o,h)$ and $\mathcal{J}_{reg,o}(o)$ in (19) are convex though not in the strict sense. Although the convergence of the algorithm to the optimal solution is theoretically difficult to prove, numerical experiments indicate that the global procedure does converge when the initialization is carried out as described above.

A delicate situation is in the choice of the regularization parameter λ_o; too small values yield overly oscillatory estimates owing to noise or discontinuities, while

too large values yield overly smooth estimates. The selection or estimation of the regularization parameter is thus a critical issue on which there have been several proposed approaches [45]. However, we are looking for a simple technique that could be combined with the AM algorithm and also fits well with the Bayesian framework. The difficulty in performing marginalization with respect to λ_o is that the partition function is not easily computed. An approach to circumvent this problem is by approximating the partition function Z_{new,λ_o} as $\lambda_o^{-N_x N_y N_z}$ [46]. By assuming a uniform hyperprior on λ_o and maximizing (19) with respect to λ_o leads to the optimal λ_o at iteration $(n+1)$ as, $\hat{\lambda}_o^{(n+1)} = (N_x N_y N_z)/\sum_{\mathbf{x}\in\Omega} |\nabla \hat{o}^{(n)}(\mathbf{x})|$.

begin

Input: Observed volume $i \in \mathbb{N}^3$.

Data: Initial parameters $\hat{\boldsymbol{\theta}}^{(0)}$ (Section 4.A), convergence criterion ϵ.

Output: Deconvolved volume $\hat{o} \in \mathbb{N}^3$, PSF parameters $\hat{\boldsymbol{\theta}} \in \Theta \subset \mathbb{R}^2_+$.

Initialization: $n \leftarrow 0$, $\hat{o}^{(n)}(\mathbf{x}) \leftarrow \texttt{Mean}(i(\mathbf{x}))$, $\hat{h}^{(n)}(\mathbf{x}) \leftarrow h(\mathbf{x}; \hat{\boldsymbol{\theta}}^{(n)})$ (22).

Estimate the background term $\hat{b}$ from the image histogram (Section 3.A).

while $|\hat{\boldsymbol{\theta}}^{(n)} - \hat{\boldsymbol{\theta}}^{(n-1)}|/\hat{\boldsymbol{\theta}}^{(n)} \geq \epsilon$ **do**

Hyperparameter λ_o estimation: $\hat{\lambda}_o^{(n)} \leftarrow 1/\texttt{Mean}(|\nabla \hat{o}^n(\mathbf{x})|)$.

Using the *minmod scheme* [29], calculate $div(\nabla \hat{o}^n(\mathbf{x})/|\nabla \hat{o}^n(\mathbf{x})|)$.

Deconvolution: Calculate $\hat{o}^{(n+1)}$ from (21).

Projection Operation: Scale $\hat{o}^{(n+1)}$ for preserving the flux (3.A.2).

Parameter estimation: Calculate $\hat{\boldsymbol{\theta}}^{(n+1)}$ from (24), (25).

Assign: $\hat{h}^{(n+1)}(\mathbf{x}) \leftarrow h(\mathbf{x}; \hat{\boldsymbol{\theta}}^{(n+1)})$ and $n \leftarrow (n+1)$.

end

end

Algorithm 1: Schema for the proposed blind deconvolution algorithm.

4.B. Numerical experiments

For the numerical experiments in Fig. 3, we have used a 3-D simulated test object of dimension $128 \times 128 \times 64$, with XY and Z pixel sizes of $20nm$ and $50nm$ respectively. The observed data was then generated by using an analytical model of the

microscope PSF (3) (with a pinhole diameter of 1 AU), and the noise is modeled as Poisson statistics (see Fig. 3(b); PSNR: 16.77dB). The results of the AM algorithm are illustrated in Fig. 3(c), (d) respectively. The stopping threshold ϵ between two successive iterations was fixed as 10^{-4}. Fig. 4 shows the reduction in the cost function with iterations of the GD algorithm and the approach of the estimated lateral spread parameter σ_r to the stable value given the estimate of the object. The quality of the restoration can be assessed by comparing with the original synthetic object using the *I-divergence* or generalized Kullback distance [29]. For the AM algorithm, when the stopping criterion ϵ was reached, the final *I-divergence* between o and $\hat{o}$ was 1.4334 (as opposed to 5.55 between o and i). Fig. 5(a) compares the estimated 3-D PSF with the analytically modeled [11] PSF and the best 3-D Gaussian fit (in the LSS sense) for the analytical model. The PSFs are shown along one direction of an off-central lateral plane, and a section of the plot can be viewed as an inset. The maximum of the residual error between the estimate and the true PSF is displayed on a logarithmic contrast stretch in Fig. 5(b). Although the Gaussian model does not capture the ringing side lobes as is evident from the residue, the RSE was found to be $< 0.07\%$.

4.C. Experiments on Real Data

4.C.1. Imaging Setup and sample description

The Zeiss LSM 510 confocal microscope is mounted on a motorized inverted stand (Zeiss Axiovert 200M) and is equipped with an *ArKr* excitation laser of wave-

length of 488nm. The Band Pass (BP) filter transmits emitted light within the band $505 - 550$nm.

The specimen that was chosen for the first experiment is an embryo of the *Drosophila Melanogaster* (see Fig. 6(a)). It was mounted and tagged with the Green Fluorescence Protein (GFP). This preparation is used for studying the sealing of the epithelial sheets (*Dorsal Closure*) midway during the *embryogenesis*. The objective lens is a Plan-Neofluar with 40X magnification having a NA of 1.3 and immersed in oil (Immersol$^{\text{TM}}$518F, Zeiss, refractive index $\mu = 1.518$). The pinhole size was 67μm. The images (© Institute of Signaling, Development Biology & Cancer, Nice UMR6543/CNRS/UNSA) were acquired with a XY pixel size of 50nm and a Z step size of 170nm, and the size of the volume imaged is $25.59 \times 25.59 \times 2.55\mu$m.

The second set of images (© INRA, Sophia-Antipolis) are the root apex of the plant *Arabidopsis Thaliana* immersed in water (see Fig. 7). The dissected roots of the *Arabidopsis Thaliana* plant were directly put on a microscope slide in approximately $100\mu l$ of water and this was then gently covered with a coverslip. This simple set up works very well when the image acquisition recording times are not too long (about 30 minutes). The microscope specifications are the same as that used for acquiring the first data set but the objective is a C-Apochromat water immersion lens with 63X magnification, 1.2 NA. The lateral pixel dimensions are 113nm and the Z step is 438nm. The pinhole was fixed at 110μm. This preparation was used to study *Nematode* infection at the center of the root in the vascular tissue.

4.C.2. Deconvolution Results

A rendered sub-volume of the observed and restored data for the *Drosophila Melanogaster* is shown in Fig. 6. The deconvolution algorithm was stopped when the difference between subsequent estimates was lower than $\epsilon = 0.002$. The AM algorithm converged after 40 iterations of the joint RL-TV and GD algorithm. The PSF parameters were initialized to 300nm and 600nm for the lateral and the axial case respectively, and the GD algorithm estimated them to be 257.9 and 477.9nm [47]. These are larger (by about 16% and 14.5% for the lateral and the axial case respectively) than their corresponding theoretically calculated values [5]. These results are fully in line with also an experimental study performed earlier [48] with sub-resolution beads which indicated a large deviation between theoretical aberration-free PSF models and empirically determined PSFs.

Fig. 8(a) shows a rendered sub-volume (as indicated in Fig. 7) of the observed root-apex and the corresponding restored result is shown in Fig. 8(b). It is evident from these results that the *microtubules* (as identified by their specific binding proteins-Microtubules binding domain (MBD)) are much easily discerned in the restoration than in the original data. It was verified from the experiments on synthetic data [47] that the proposed algorithm can not only estimate the actual PSF, but also provide much better deconvolution result [49] in comparison to theoretical microscope PSF's (generated using the microscope settings). Validation is very important as in some

situations artifacts might arise in the restored image. These artifacts would be hard to distinguish from biological structures unless some knowledge about the true image is available. However, the results on real data are difficult to be validated unless a higher resolution image of the same sample is available. Hence, we tested our deconvolution algorithm on images of spherical fluorescent shells (see [36]) whose thickness was measured after deconvolution and found closer to the true value specified by Molecular Probes®.

5. Conclusions and future work

In this paper we have proposed and validated an "*Alternate Minimization (AM)*" algorithm for the joint estimation of the microscope PSF and the specimen source distribution for a CLSM. We choose the RL algorithm for the deconvolution process as it is best suited for the Poisson data, and TV as the regularization model. A separable 3-D Gaussian model best describes the diffraction-limited confocal PSF, and is chosen as the a priori model for the PSF. We are able to achieve blind deconvolution by constraining the solution of the object and the PSF to different spaces. The PSF approximation that is given in this paper is currently relevant to imaging thin samples. However, it could also be extended to encompass any PSF that can be decomposed in a similar manner. We have experimented on simulated and real data, and the method gives very good deconvolution results and a PSF estimation close to the true value [29, 47]. However, it should be noted that, all of the out-of-focus light cannot

be rejected and some noticeable haze and axial smearing remains in the images. This could be improved by adding a Gamma prior on the PSF parameters.

Acknowledgement

This research was partially funded by the P2R Franco-Israeli Collaborative Research Program. The authors gratefully acknowledge Dr. Caroline Chaux (Université Paris-Est, France) and Prof. Arie Feuer (Technion, Israel) for several interesting discussions. We would also like to thank INRIA for supporting the Ph.D. of the first author through a CORDI fellowship and CNRS for supporting the Ph.D. of the second author. Additionally, our sincere gratitude goes to Dr. Stéphane Noselli, Mrs. Fanny Serman (Ph.D. candidate) from the Institute of Signaling, Development Biology & Cancer UMR 6543/CNRS/UNSA and Dr. Gilbert Engler (INRA Sophia-Antipolis, France) for painstakingly preparing the images presented in Section 4.C, and for the useful comments on their validation.

References

1. J. B. Pawley, editor. *Handbook of biological confocal microscopy.* Springer, 3^{rd} edition, 2006.
2. M. Born and E. Wolf. *Principles of Optics.* Cambridge Press, Cambridge, 1999.
3. D. A. Agard. Optical sectioning microscopy: Cellular architecture in three dimensions. *Ann. Rev. Biophys. Bioeng.*, 13:191–219, 1984.

4. R. Hudson, J. N. Aarsvold, C.-T. Chen, J. Chen, P. Davies, T. Disz, I. Foster, M. Griem, M. K. Kwong, and B. Lin. Optical microscopy system for 3D dynamic imaging. In C. J. Cogswell, G. S. Kino, and T. Wilson, editors, *Proc. SPIE*, volume 2655, pages 187–198, 1996.

5. B. Zhang, J. Zerubia, and J. C. Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.*, 46(10):1819–1829, 2007.

6. J. G. McNally, C. Preza, J.-Á. Conchello, and L. J. Thomas Jr. Artifacts in computational optical-sectioning microscopy. *J. Opt. Soc. Am. A*, 11:1056–1067, March 1994.

7. J. W. Shaevitz and D. A. Fletcher. Enhanced three-dimensional deconvolution microscopy using a measured depth-varying point-spread function. *J. Opt. Soc. Am. A*, 24(9):2622–2627, 2007.

8. P. J. Shaw and D. J. Rawlins. The point-spread function of a confocal microscope: its measurement and use in deconvolution of 3-D data. *Journal of Microscopy*, 163(2):151–165, 1991.

9. P. J. Shaw. Deconvolution in 3-D optical microscopy. *The Histochemical Journal*, 26(9):687–694, Sept. 1994.

10. A. Dieterlen, M. Debailleul, A. De Meyer, B. Simon, V. Georges, B. Colicchio, O. Haeberle, and V. Lauer. Recent advances in 3-D fluorescence microscopy:

tomography as a source of information. In M. Kujawinska and O. V. Angelsky, editors, *Proc. SPIE*, volume 7008, pages 70080S1–8. SPIE, 2008.

11. P. A. Stokseth. Properties of a Defocused Optical System. *J. Opt. Soc. Am. A*, 59:1314–1321, October 1969.

12. S.F. Gibson and F. Lanni. Diffraction by a circular aperture as a model for three-dimensional optical microscopy. *J. Opt. Soc. Am. A*, A6:1357–1367, 1989.

13. O. V. Michailovich and D. R. Adam. Deconvolution of medical images from microscopic to whole body images. In P. Campisi and K. Egiazarian, editors, *Blind Image Deconvolution: Theory and Applications*, pages 169–237. CRC Press, 2007.

14. T. J. Holmes. Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach. *J. Opt. Soc. Am. A*, 9:1052–1061, July 1992.

15. J. Markham and J.-A. Conchello. Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur. *J. Opt. Soc. Am. A*, 16(10):2377–2391, 1999.

16. E. F. Y. Hom, F. Marchis, T. K. Lee, S. Haase, D. A. Agard, and J. W. Sedat. AIDA: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data. *J. Opt. Soc. Am. A*, 24(6):1580–1600, 2007.

17. L. Mandel. Sub-Poissonian photon statistics in resonance fluorescence. *Opt. Lett.*, 4:205–207, July 1979.

18. D. A. Agard, Y. Hiraoka, P. Shaw, and J. W. Sedat. Fluorescence microscopy in three dimensions. *Methods Cell Biology*, 30:353–377, 1989.

19. G. B. Avinash. Data-driven, simultaneous blur and image restoration in 3-D fluorescence microscopy. *Journal of Microscopy*, 183(2):145–157, August 1996.

20. P. Pankajakshan, L. Blanc-Féraud, Z. Kam, and J. Zerubia. Point-Spread Function retrieval in Fluorescence Microscopy. In *Proceedings of IEEE International Symposium on Biomedical Imaging*, July 2009. Accepted for publication.

21. B. M. Hanser, M. G. Gustafsson, D. A. Agard, and J. W. Sedat. Phase retrieval for high-numerical-aperture optical systems. *Opt. Lett.*, 28(10):801–803, May 2003.

22. C. J. R. Sheppard and C. J. Cogswell. Three-dimensional image formation in confocal microscopy. *Journal of Microscopy*, 159(2):179–194, 1990.

23. A. Erhardt, G. Zinser, D. Komitowski, and J. Bille. Reconstructing 3-D light-microscopic images by digital image processing. *Appl. Opt.*, 24:194–200, 1985.

24. W.A. Carrington, K. E. Fogarty, and F. S. Fay. *3D Fluorescence Imaging of Single Cells Using Image Restoration*, pages 53–72. Wiley-Liss, 1990.

25. C. Preza, M. I. Miller, L. J. Thomas Jr., and J. G. McNally. Regularized linear method for reconstruction of three-dimensional microscopic objects from optical sections. *J. Opt. Soc. Am. A*, 9(2):219–228, 1992.

26. T. Tommasi, A. Diaspro, and B. Bianco. 3-D reconstruction in optical microscopy by a frequency-domain approach. *Signal Processing*, 32(3):357–366, 1993.

27. T. J. Holmes. Maximum-likelihood image restoration adapted for noncoherent optical imaging. *J. Opt. Soc. Am. A*, 5:666–673, May 1988.

28. N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, P. Roux, J.C. Olivo-Marin, and J. Zerubia. Richardson-Lucy Algorithm with Total Variation Regularization for 3D Confocal Microscope Deconvolution. *Microscopy Research Technique*, 69:260–266, 2006.

29. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.C. Olivo-Marin, and J. Zerubia. Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy (CLSM)-Proof of Concept. Research Report 6493, INRIA Sophia-Antipolis, France, March 2008.

30. G. M. P. Van Kempen, L. J. Van Vliet, P. J. Verveer, and H. T. M. Van Der Voort. A quantitative comparison of image restoration methods for confocal microscopy. *Journal of Microscopy*, 12:354–365, March 1997.

31. A. Dieterlen, C. Xu, O. Haeberle, N. Hueber, R. Malfara, B. Colicchio, and S. Jacquey. Identification and restoration in 3D fluorescence microscopy. In Oleg V. Angelsky, editor, *Proc. SPIE*, volume 5477, pages 105–113. SPIE, 2004.

32. G. Demoment. Image reconstruction and restoration: overview of common estimation structures and problems. *IEEE Trans. Acoustics, Speech and Signal Processing*, 37(12):2024–2036, Dec 1989.

33. A. N. Tikhonov and V. A. Arsenin. *Solution of Ill-posed Problems.* Winston &

Sons, 1977.

34. K. Miller. Least Squares Methods for Ill-Posed Problems with a Prescribed Bound. *SIAM Journal on Mathematical Analysis*, 1(1):52–74, 1970.

35. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D.*, 60:259–268, 1992.

36. N. Dey, L. Blanc-Féraud, C. Zimmer, P. Roux, Z. Kam, and J.C. Olivo-Marin. 3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization. Research Report 5272, INRIA, France, July 2004.

37. L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astron. J.*, 79:745–754, 1974.

38. W. H. Richardson. Bayesian-Based Iterative Method of Image Restoration. *J. Opt. Soc. Am. A*, 62(1):55–59, January 1972.

39. A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society B*, 39(1):1–38, 1977.

40. M. Jiang and G. Wang. Development of blind image deconvolution and its applications. *Journal of X-Ray Science and Technology*, 11:13–19, 2003.

41. T. F. Chan and C.-K. Wong. Total variation blind deconvolution. *IEEE Trans. Image Processing*, 7(3):370–375, Mar 1998.

42. L. Bar, N. A. Sochen, and N. Kiryati. Variational Pairing of Image Segmenta-

tion and Blind Restoration. In T. Pajdla and J. Matas, editors, *Proc. ECCV*, volume II, pages 166–177. Springer, May 2004.

43. A. Santos and I. T. Young. Model-Based Resolution: Applying the Theory in Quantitative Microscopy. *Appl. Opt.*, 39(17):2948–2958, 2000.

44. K. E. Atkinson. *An introduction to Numerical Analysis*. John Wiley and Sons, 2nd edition, 1989.

45. A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Hyperparameter estimation for satellite image restoration using a MCMC Maximum Likelihood method. *Pattern Recognition*, 35(2):341–352, 2002.

46. A. Mohammad-Djafari. A full Bayesian approach for inverse problems. In K. Hanson and R. N. Silver, editors, *Maximum entropy and Bayesian methods*, volume 79, pages 135–143. Kluwer Academic Publisher, 1996.

47. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.C. Olivo-Marin, and J. Zerubia. Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy. In *Proceedings of IEEE International Conference of EMBS*, pages 6531–6534, Lyon, France, August 2007.

48. M. de Moraes Marim, B. Zhang, J.-C. Olivo-Marin, and C. Zimmer. Improving single particle localization with an empirically calibrated Gaussian kernel. In *Proceedings of IEEE International Symposium on Biomedical Imaging*, pages 1003–1006, May 2008.

49. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.C. Olivo-Marin, and J. Zerubia. Blind deconvolution for diffraction-limited fluorescence microscopy. In *Proceedings of IEEE International Symposium on Biomedical Imaging*, pages 740–743, Paris, May 2008.

List of Figure Captions

Fig. 1. The MRF over a 6 member neighborhood $\eta_{\mathbf{x}}$.

Fig. 2. Variation of the energy function $\mathcal{J}(o, \boldsymbol{\theta}|i)$ with respect to (a) lateral (σ_r) and (b) with axial PSF parameter (σ_z). For this experiment, the true object o is known and the observation is generated using a known 3-D Gaussian model. The axial PSF parameter σ_z is varied by a factor $\pm\epsilon$ to monitor its effect on the estimated parameter σ_r and vice versa. $\sigma_{(\cdot,true)}$ is the true parameter value.

Fig. 3. 3-D (a) phantom object (with false coloring), (b) observed image blurred by the PSF model (3) and Poisson noise (PSNR: 16.77dB, I-divergence: 5.55), (c) restoration after RL+TV deconvolution with the estimated PSF (I-divergence: 1.43), (d) estimated PSF. The intensities of the object, observation and the restoration are on a linear scale while the PSF is on a logarithmic scale.

Fig. 4. Convergence of the cost function and lateral parameter by the GD method (when the original object is known). The Y axis is left-scaled for the cost function $\mathcal{J}(\hat{\boldsymbol{\theta}}, \hat{o}|i)$ and right-scaled for the PSF parameter respectively.

Fig. 5. (a) The full model (dash), estimated (continuous) and the best Gaussian fit (dash-dot) PSFs are displayed for one direction (off-central plane); the inset shows a section of the plot, (b) X-Z projection of the residual (RSE $< 0.07\%$) between the estimated and full PSF model is displayed on a log scale.

Fig. 6. (a) Rendered sub-volume of the original specimen (© Institute of

Signaling, Developmental Biology & Cancer UMR6543/CNRS/UNSA), and (b) restored image (© Ariana-INRIA/I3S). The intensity is scaled between [0 130] for display.

Fig. 7. Observed root apex of an *Arabidopsis Thaliana* with a volume $146.448\mu\text{m} \times 146.448\mu\text{m} \times 30.222\mu\text{m}$ (© INRA). The sub-volume chosen for restoration is emphasized.

Fig. 8. Rendered sub-volume of the (a) observed image slices in Fig. 7 (© INRA) and (b) volume rendering of the restored image slices (© Ariana-INRIA/I3S). $\epsilon = 0.0001$.

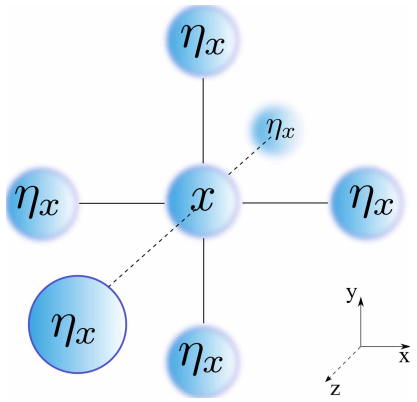
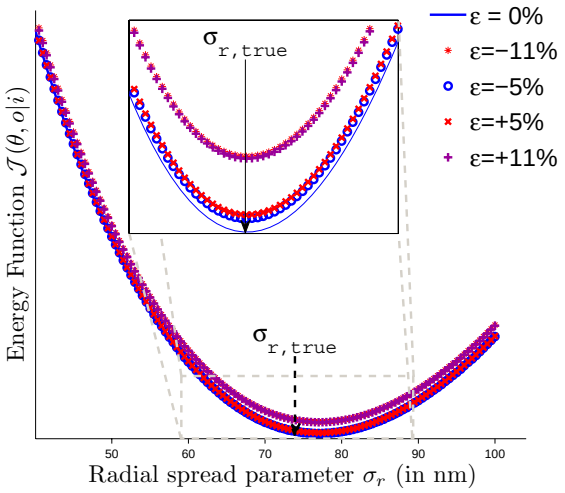
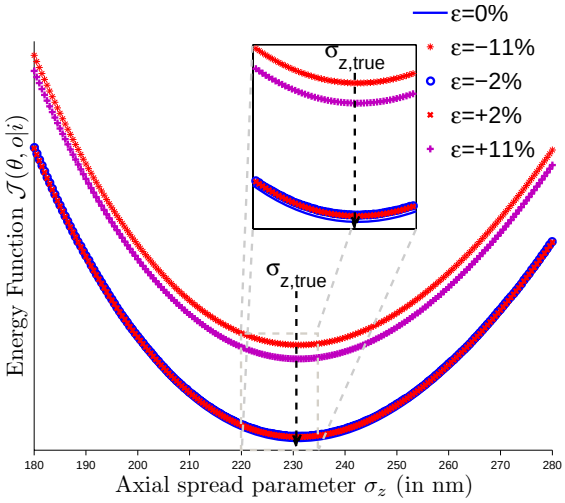


Fig. 1. The MRF over a 6 member neighborhood $\eta_{\mathbf{x}}$. fig1.eps.



(a)



(b)

Fig. 2. Variation of the energy function $\mathcal{J}(o, \boldsymbol{\theta}|i)$ with respect to (a) lateral (σ_r) and (b) with axial PSF parameter (σ_z). For this experiment, the true object o is known and the observation is generated using a known 3-D Gaussian model. The axial PSF parameter σ_z is varied by a factor $\pm\epsilon$ to monitor its effect on the estimated parameter σ_r and vice versa. $\sigma_{(\cdot,true)}$ is the true parameter value. fig2a.eps fig2b.eps.

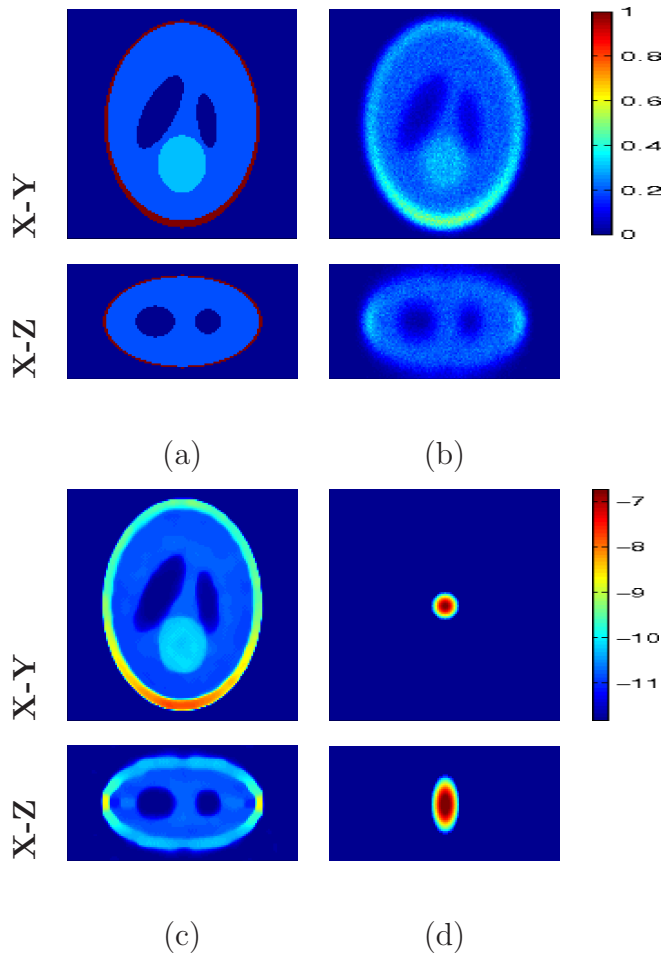


Fig. 3. 3-D (a) phantom object (with false coloring), (b) observed image blurred by the PSF model (3) and Poisson noise (PSNR: 16.77dB, I-divergence: 5.55), (c) restoration after RL+TV deconvolution with the estimated PSF (I-divergence: 1.43), (d) estimated PSF. The intensities of the object, observation and the restoration are on a linear scale while the PSF is on a logarithmic scale. fig3axy.eps fig3bxy.eps fig3scale.eps fig3axz.eps fig3bxz.eps fig3cxy.eps fig3dxy.eps fig3scalepsf.eps fig3cxz.eps fig3dxz.eps.

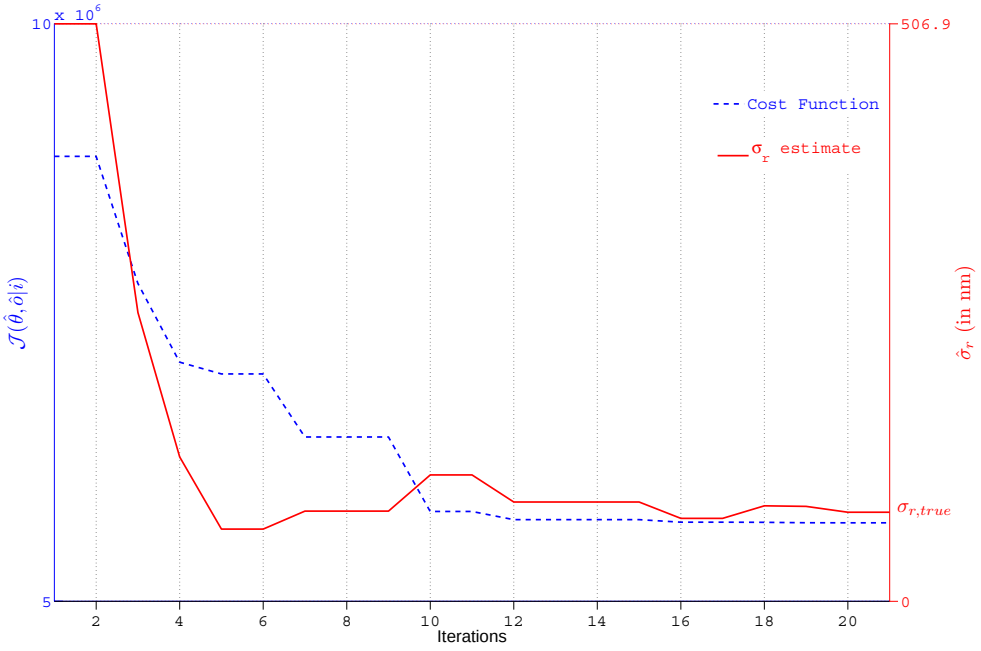
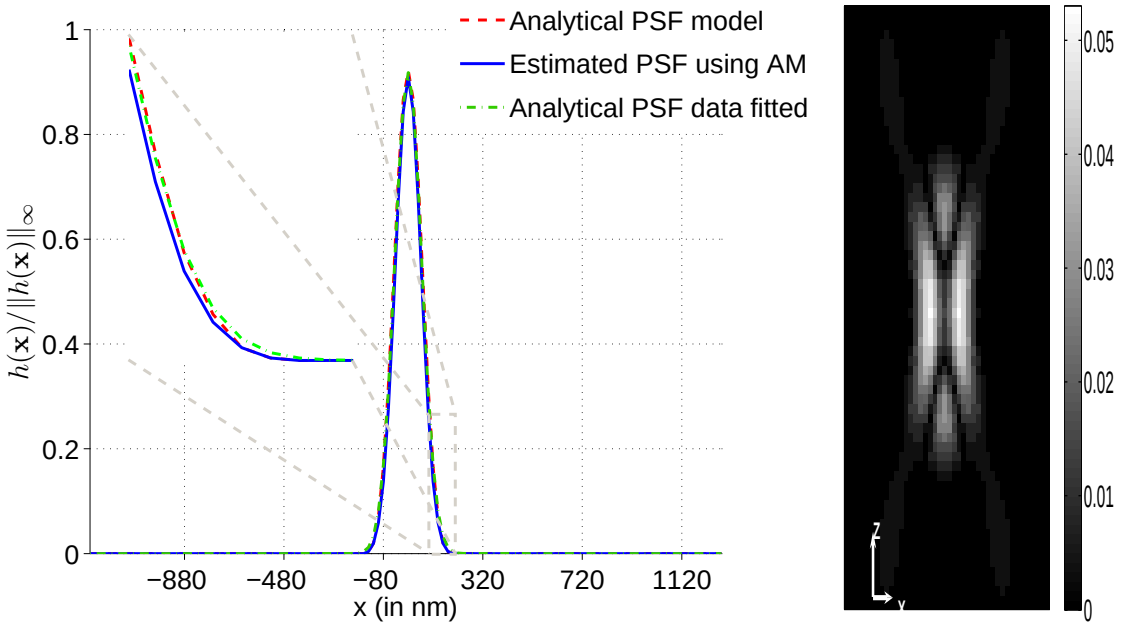
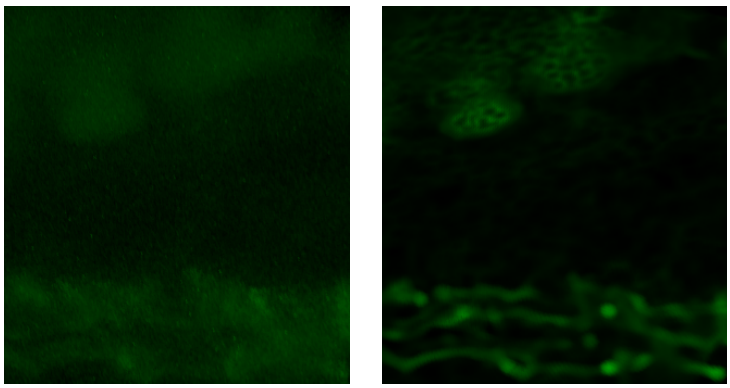


Fig. 4. Convergence of the cost function and lateral parameter by the GD method (when the original object is known). The Y axis is left-scaled for the cost function $\mathcal{J}(\hat{\boldsymbol{\theta}}, \hat{o}|i)$ and right-scaled for the PSF parameter respectively. fig4.eps.



(a) (b)

Fig. 5. (a) The full model (dash), estimated (continuous) and the best Gaussian fit (dash-dot) PSFs are displayed for one direction (off-central plane); the inset shows a section of the plot, (b) X-Z projection of the residual (RSE $< 0.07\%$) between the estimated and full PSF model is displayed on a log scale. fig5a.eps fig5b.eps.



(a) (b)

Fig. 6. (a) Rendered sub-volume of the original specimen (© Institute of Signaling, Developmental Biology & Cancer UMR6543/CNRS/UNSA), and (b) restored image (© Ariana-INRIA/I3S). The intensity is scaled between [0 130] for display. fig6a.eps fig6b.eps.

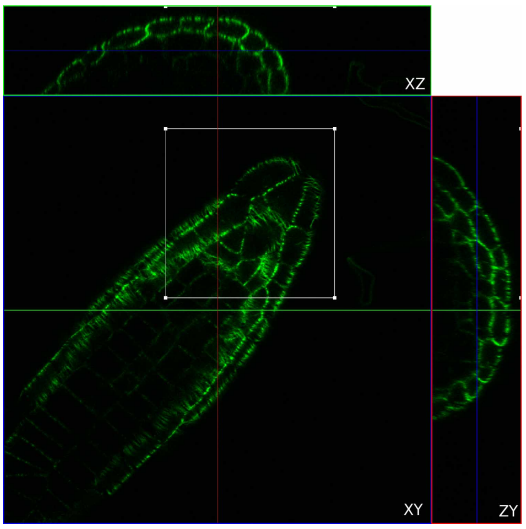
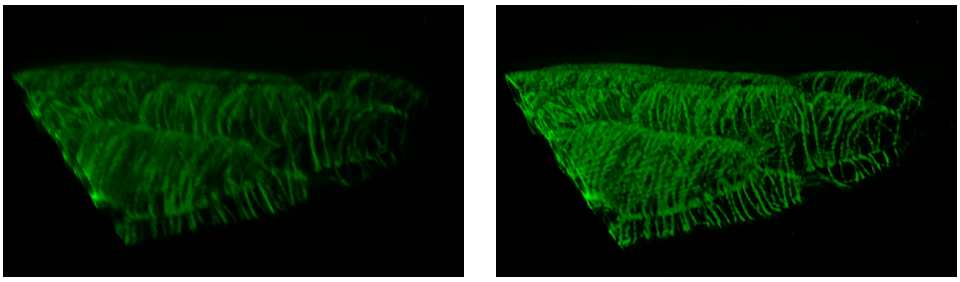


Fig. 7. Observed root apex of an *Arabidopsis Thaliana* with a volume 146.448μm $\times$ 146.448μm $\times$ 30.222μm (© INRA). The sub-volume chosen for restoration is emphasized. fig7.eps.



(a) (b)

Fig. 8. Rendered sub-volume of the (a) observed image slices in Fig. 7 (© INRA) and (b) volume rendering of the restored image slices (© Ariana-INRIA/I3S). $\epsilon = 0.0001$. fig8a.eps fig8b.eps.